



Context-Based Scalable Coding and Representation of High Resolution Art Pictures for Remote Data Access

Marie Babel, Olivier Déforbes, Laurent Bédard, Jean Motsch

► To cite this version:

Marie Babel, Olivier Déforbes, Laurent Bédard, Jean Motsch. Context-Based Scalable Coding and Representation of High Resolution Art Pictures for Remote Data Access. IEEE International Conference on Multimedia and Expo, ICME'07, Jul 2007, Beijing, China. pp.460-463. hal-00163329

HAL Id: hal-00163329

<https://hal.science/hal-00163329>

Submitted on 17 Jul 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONTEXT-BASED SCALABLE CODING AND REPRESENTATION OF HIGH RESOLUTION ART PICTURES FOR REMOTE DATA ACCESS

Marie Babel ^a, Olivier Déforges ^a, Laurent Bédet ^a, Jean Motsch ^b

IETR UMR CNRS 6164, Image and Remote Sensing Group, FRANCE

^a INSA of Rennes, 35043 RENNES. Contact: {mbabel,odeforge,lbedat,}@insa-rennes.fr

^b LESTP/CREC St-Cyr, 56381 Guer. Contact: jean.motsch@st-cyr.terre.defense.gouv.fr

ABSTRACT

EROS is the largest database in the world of high resolution art pictures. The TSAR project is designed to open it in a secure, efficient and user-friendly way that involves cryptography and watermarking as well as compression and region-level representation abilities. This paper more particularly addresses the two last points. The LAR codec is first presented as a suitable solution for picture encoding with compression ranging from highly lossy to lossless. Then, we detail the concept of self-extracting region representation, which consists of performing a segmentation process at both the coder and decoder from a highly compressed image, and later locally enhancing the image in a region of interest. The overall scheme provides an efficient, consistent solution for advanced data browsing.

1. INTRODUCTION

Museums are supposed to undertake at least two essential missions [1]. Firstly, they have to preserve their huge number of items and save them from damage. At the same time, museums play an active role in the spread of cultural knowledge and this educational objective leads them to widely communicate these materials. However, these two missions are somewhat contrary in nature because handling art items inevitably causes damage. To solve this major problem, museum research centers have introduced the “digital museum” concept [2]: digital versions of the original art items are collected in a database on a server accessible via the Internet.

For example, the National Gallery in London, the Tokyo University Digital Museum (through the Digital Museum 2000 project [3]) or the Chinese University Museum Grid [1] provide public access to their databases. However, users can only download low-resolution images. The application has been actually designed to prevent illegal copies of digitized data and the best way to achieve this consists basically of not transmitting high-resolution images. High-resolution information is therefore stored separately and reserved for a limited number of people.

In France, the C2RMF laboratory, connected to the Louvre museum, has digitized more than 300,000 documents taken from French museums, in high resolution (up to 20,000 × 30,000 pixels). The resulting EROS database [4] is for the moment only accessible to researchers whose work is connected with the C2RMF. In order to widely open the database, two problems have to be resolved. First, as the amount of data is huge, an efficient compression method has to be applied. In particular, for research purposes, this compression scheme has to provide a lossless solution. Secondly, the application must support interactive actions that can make browsing easier for users. Then, the coding method has to integrate advanced functionalities such as full scalability and Region Of Interest (ROI) coding.

As far as the French TSAR project is concerned, the LAR image coding method has been chosen for this purpose. The objective is to design an art image database accessible through a client-server process that includes and combines a hierarchical description of images and a hierarchical secured access.

This article is organized as follows. In section 2, the TSAR project is described as the context of our study. The LAR method and its scalable lossless extensions are presented in section 3. Section 4 is devoted to an efficient ROI encoding solution involving an original segmentation process. The principles of the corresponding client-server application are developed in section 5.

2. THE TSAR PROJECT

The TSAR project (Secure Transmission of high-Resolution Art images) is supported by the French National Research Agency. Five laboratories are involved, namely the C2RMF (Louvre, Paris), IETR (Rennes), IRCCyN (Nantes), LIRMM (Montpellier) and LIS (Grenoble). The C2RMF has developed the EROS (European Research Open System) database storing digital art documentation[4]. Till now, images have been stored in pyramidal TIFF format involving a bit overhead of 33%. As the amount of data is tremendous, this solution is no longer acceptable. The idea is to integrate another scalable coding solution able to achieve a high lossy and lossless compression ratio. A second area of research concerns the secure access of images. The objective is to build hierarchical secure protection based on the rights of various user groups.

3. LAR METHOD FOR IMAGE COMPRESSION

3.1. Basic LAR codec

The LAR method has been initially introduced for lossy image coding [5]. It is based on the assumption that an image can be represented as layers of basic information and local texture. Thus, the overall scheme of this approach consists of two scalable layers (see Fig. 1): a first one to encode an image at low bit rates, and a second one for visual quality enhancement at medium/high bit rates.

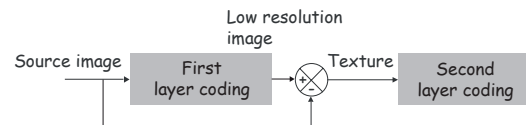


Fig. 1. Basic description of the LAR coder: two-layer coding

The method relies first on Quadtree partitioning of the image.

Homogeneity is provided by a local gradient estimation that fixes small blocks onto contours and large ones onto flat areas. The Quadtree partition is denoted $QP^{[N_{max} \dots N_{min}]}$ with all square blocks having a size equal to a power of two, and N_{max} and N_{min} representing respectively the maximum and minimum authorized block sizes. The first layer representation consists of rebuilding each block by its mean luminance value, encoded using a DPCM approach. The low bit-rate rebuilt image is visually acceptable thanks to the Quadtree partition, that accommodates the variable block-size as a function of the original image context.

To encode texture, the second layer performs a variable block-size DCT transform to fit the $QP^{[N_{max} \dots N_{min}]}$ partition. Scalability can then be content-dependent: the enhancement of contours presupposes that only the smallest blocks are processed. Subjective quality tests have been carried out between JPEG, JPEG2000 and LAR codecs: LAR outperforms other techniques for nearly all the test sequences [5].

3.2. Multiresolution LAR codec

3.2.1. The overall scheme

To both increase scalability capacity and address lossless compression, we have proposed multiresolution extensions of the basic LAR called Interleaved S+P [6] and RWHaT+P [7]. The overall approach used in these two techniques is identical; the only difference lies in the decomposition step. To fit the Quadtree partition, diadic decomposition is carried out. The first and second layers of the basic LAR are replaced by two successive pyramidal descent processes, but the image representation content is preserved: the first decomposition reconstructs the low-resolution image (LAR-image) while the second one processes the local texture information. Thus, the first pyramid pass performs a conditional decomposition in accordance with the Quadtree partition. We consider a pyramid between

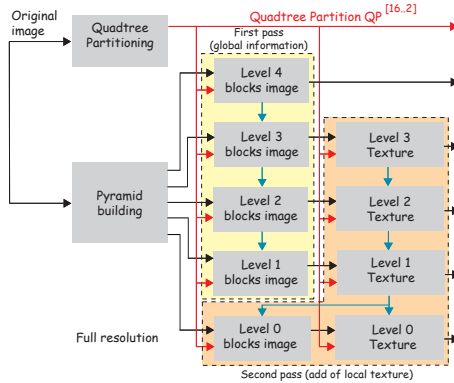


Fig. 2. Pyramidal conditional decomposition of the LAR image in accordance with the Quadtree partition $QP^{[16..2]}$

levels 0 (full resolution) and L_{max} . A pyramid representation involving a diadic decomposition is generally associated with a multi-level Quadtree partition $QP^{[2^{L_{max}} \dots 2^l]}$ where level l of the pyramid also specifies the finest resolution. More generally, we consider that the Quadtree partition $QP^{[N_{max} \dots N_{min}]}$ defines only allowed block sizes. An additional parameter $N_l \in [N_{max} \dots N_{min}]$ gives the upper limit of block sizes to be decomposed at level l of the pyramid. For instance, a global partition $QP^{[64 \dots 2]}$ leads to the encoding of the representation only from sizes 64 to 2, while $N_0 = 4$ means that blocks of sizes 4 and 2 will be decomposed at level 0.

The second pyramid pass processes the local texture information. All blocks at the current level which have not been encoded during the first pass are decomposed by the refinement layer.

Figure 2 gives the global scheme of the multiresolution LAR coder, for a partition $QP^{[16..2]}$, $N_l = 2^l$ for $1 \leq l \leq L_{max}$ and $N_0 < 2$.

The use of Quadtree decomposition has several advantages:

1. it doubles the number of decomposition levels ($2 \times L_{max}$), thereby increasing scalability,
2. good quality images are available at low bit rates,
3. the approach acts as “objective context modeling”, removing the correlation of error prediction laws between high entropy features during the first pass, and low entropy features during the final pass.

3.2.2. The decomposition method

To achieve lossless compression, the decomposition method involved in the pyramid descent process has to be reversible. In [6] we proposed a modified version of the S+P transform [8] with advanced prediction. Recently, the RWHaT transform has been introduced [7]. RWHaT means “Reversible Walsh Hadamard Transform (WHT)”, and we present a reversible transformation based on the standard WHT for 2×2 blocks, based on original rounding operations. We also add a prediction step using inter and intra levels in the pyramid, but it can also act as an interpolation stage. A unified solution for prediction and interpolation can be a worthwhile feature for the progressive rebuilding of images: an image can first be interpolated from the previous level, then the pixel values can be used for prediction, without further processing.

The proposed approach is efficient for lossy compression using the quantization step as well as for lossless coding. In this latter context, we show that it outperforms state-of-the-art methods for natural images. The same experiments have been carried out on a set of art pictures. A comparison of results with CALIC [9] (which is not scalable) is reported in table 1.

	Image	Size	Raw	CALIC	RWHaT
	Louvre 1	1200× 1840	7.22	4.51	4.24
	Louvre 2	2368× 3536	5.81	3.87	3.82
	Louvre 3	1392× 1120	7.72	3.29	3.25
	Louvre 4	3848× 3382	6.52	3.73	3.57

Table 1. Comparison for lossless coding between the proposed approach and the state-of-the-art method. First-order entropy (bit/pixels).

4. IMAGE REPRESENTATION WITH THE SELF-EXTRACTING REGION PROCESS

4.1. Principles

The region-based approach provides advanced functionalities such as interaction between objects and regions, or scene composition.

Another important advantage is the ability to accurately increase image quality only within a region of interest. Nevertheless, region-based techniques generally present some significant limitations :

1. Shape description produces an information overhead, which can be fairly significant at low bit rates.
2. Region-based methods mainly preserve the "shape" component and often neglect the "content" component. Consequently, for a given representation, an encoded shape becomes independent of its content.
3. Common region-based coding schemes merely allow the encoder to define region representation; the decoder does not have any decision-making function. This type of approach cannot, therefore, be used for image database browsing, where the operator would define and select his own ROI.

To overcome this problem, we have designed an original region-based representation technique adapted to the LAR coding method. A first solution has been proposed in [5]. To avoid the prohibitive cost of this description, a suitable solution consists of performing the segmentation directly, in both the coder and decoder, using only a low bit rate compressed image. Once the region representation is built, either the coder or decoder can select the ROI for enhancement to higher quality. This process fits perfectly with a scalable coding scheme starting with a low-quality image and progressively refining it through successive compressed bitstreams.

A segmentation can be considered on a compressed image whenever distortions introduced by the encoding stage remain limited. At low bit rates, standard methods generate degradations, in particular on contours, preventing reliable segmentation results. We suggest that the segmentation should be performed on a high compressed image at block level (first layer of LAR codec), using an iterating merging process. The block-based encoding of LAR, based on a coherent representation in terms of contours and uniform areas, avoids such damaging degradation.

The segmentation algorithm is based on the Region Adjacency Graph theory. Its leads to an hierarchical region representation. Figure 3 gives some examples of region representation at different levels: segmentation has been performed here on a low resolution image (first layer including quantization step).

4.2. ROI coding

One direct application for self-extracting region representation is found in a coding scheme with local enhancement in regions of interest. From the segmentation map simultaneously available in both coder and decoder, either device can define its own ROI as a set of regions. Thus, an ROI will simply be specified by the labels of its regions. The definition of an ROI is generally performed at the highest level of the segmentation hierarchy (limited number of regions). For an ROI composed of n regions, only n labels are required to fully describe it: this represents a very small number of bytes. The method provides both a *semi-automatic tool for ROI selection*, and *probably the best solution for its concise definition*.

Each region, and consequently each ROI, consists of a set of blocks defined in the initial Quadtree partition. The *enhancement of an ROI is straightforward as it merely requires execution of the second layer for the validated blocks, i.e. those inside the ROI*. This means that there is immediate, total compatibility between the shape of the ROI and its coding content. The ROI coding process is illustrated in figure 4. The ROI is initially selected from available regions. In this example, the ROI consists only of 7 regions, so an equal number of integer values for region labels is required to create

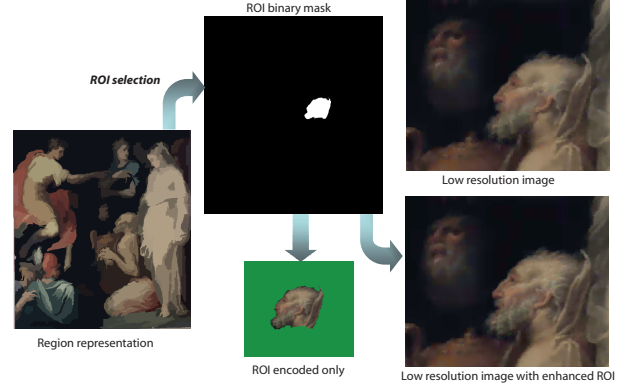


Fig. 4. ROI coding process

the binary mask of the ROI. The ROI is then used to further process the picture inside the ROI only. The enhanced ROI can be extracted alone for scene composition, or be replaced into the low-resolution image.

5. CLIENT-SERVER APPLICATION

We are currently working on a corresponding client-server application. Every client will be authorized to browse the low-resolution image database and the server application will verify the user access level for each image and ROI request. If a client application sends a request that does not match the user access level, the server application will reduce the image resolution according to access policy. The exchange protocol is depicted in figure 5.

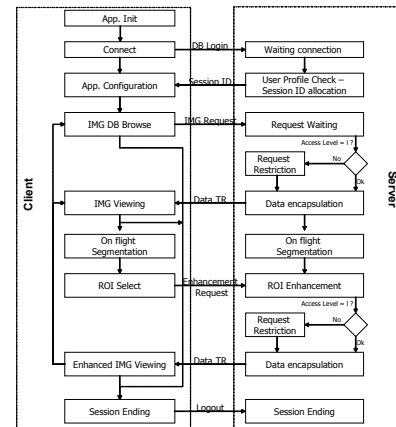


Fig. 5. Exchange protocol for client-server application

6. CONCLUSION

The opening of art picture databases represents a promising method for the spread of cultural knowledge but it demands a number of requirements in terms of reproduction rights and access control, along with fast and user-friendly browsing.

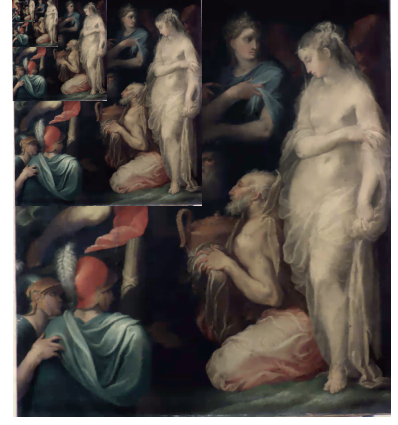
This paper has presented our contribution to the TSAR project, based on an efficient and original solution for the region-based scalable transmission of pictures. The LAR codec is a content-based compression method enabling reliable segmentation at low bit rates.



a) Original image
size : 1136×1264 , 24 bpp



b) Block-level representation with grid
58703 blocks



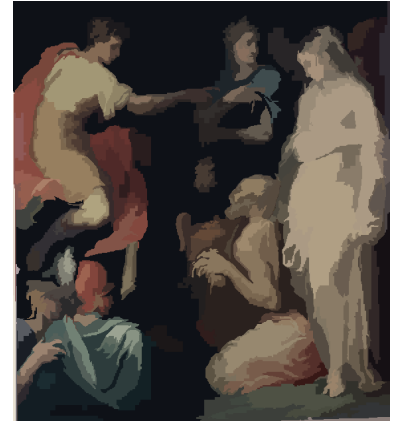
c) Highly compressed image (1st layer)
full resolution : 0.16 bpp



d) Region-level representation
Level 1 : 1812 regions



e) Region-level representation
Level 2 : 624 regions



f) Region-level representation
Level 3 : 314 regions

Fig. 3. Low resolution LAR image and hierarchical region representation on “CRT” with partition $QP^{[16...2]}$ and $N_0 < 2$

The available region-level image description can then be used to easily select an ROI and enhance its visual quality. The design of a client-server application is currently underway.

Parallel works in the TSAR project concern the security aspects (cryptography and watermarking).

Acknowledgment: This work is supported by the French National Research Agency as part of the TSAR project. All art pictures are the property of the C2RMF laboratory.

7. REFERENCES

- [1] Xiaowu Chen, Haifen Ou, Xixi Luo, Mingji Chen, Yi Zhang, Ke Hao, and Shunag Mi, “The Progress of University Digital Museum Grid,” in *Proc. IEEE International Conference on Grid and Cooperative Computing Workshops, GCCW’06*, 2006.
- [2] George F. MacDonald, “Digital Visionary,” *Museum News*, March/April 2000.
- [3] Ken Sakamura, “Digital Museum - Distributed Museum Concept for the 21st Century,” 2000.
- [4] D. Pitzalis, R. Pillay, and C. Lahanier, “A new Concept in high Resolution Internet Image Browsing,” in *10th International Conference on Electronic Publishing (ELPUB)*, June 2006.
- [5] O. Déforges, M. Babel, L. Bédar, and J. Ronsin, “Color lar codec: a color image representation and compression scheme based on local resolution adjustment and self-extraction region representation,” *IEEE Transactions on Circuits and Systems for Video Technology*, in revision, 2007.
- [6] M. Babel, O. Déforges, and J. Ronsin, “Interleaved S+P Pyramidal Decomposition with Refined Prediction Model,” in *IEEE International Conference on Image Processing, ICIP’05*, Genova, Italy, September 2005, vol. 2, pp. 750–753.
- [7] O. Déforges, M. Babel, and J. Motsch, “The RWHT+P for an improved lossless multiresolution coding,” in *European Signal Processing Conference, EUSIPCO’06*, To be published, September 2006.
- [8] A. Said and W. Pearlman, “Reversible image compression via multiresolution representation and predictive coding,” in *Visual Communication and Image Processing*. SPIE, November 1993, vol. 209, pp. 664–674.
- [9] X. Wu and N. Memon, “A Context-based, Adaptive, Lossless/Nearly-Lossless Coding Scheme for Continuous-Tone Images (CALIC),” *International Standards Organization working document, ISO/IEC SC29/WG 1/N256*, 1995.